DOCUMENT RESUME

ED 102 605 CS 201 905

AUTHOR Anderson, Scarvia B.; Dobbin, John E.
TITLE Testing and Tests: Pedagogical Versus Public Uses.
PUB DATE Nov 74
NOTE 15p.; Paper presented at the Annual Meeting of the National Council of Teachers of English (64th, New Orleans, November 28-30, 1974); Appendix A has marginal reproducibility because of type size

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Behavioral Objectives; *Educational Accountability; *Educational Testing; *Educational Theories; Elementary Secondary Education; Performance Criteria; Public Officials; Test Construction; *Testing; *Tests

ABSTRACT

Public uses of tests and testing include all those materials and practices in observation of human behavior that are intended to help administrators, school boards, legislatures, taxpayers, and others to evaluate their educational systems. Pedagogical uses of tests, on the other hand, cover all those materials and practices in observation of human behavior that are intended to help the teacher do a better job of teaching or the learner do a better job of learning, or both. The movement toward accountability affects public testing since if what legislators want students to demonstrate is reading and writing skill, then that is what the tests for public use will measure. Economic considerations dictate that tests of reading and writing skills be used to ascertain the general effectiveness of the whole organization of public education rather than that of single schools or individual teachers. Most important in the use of pedagogical tests is that they cover all the kinds of learning teachers hope to have the students attain. To accomplish this end, teachers can choose among standardized tests, teacher-made tests, and systematic observation of pupils. A good teacher who is also informed about measurement blends teaching and testing and learning so smoothly that it is often impossible to tell where one stops and another starts. (HOD)

BEST COPY AVAILABLE

National Council of Teachers of English
New Orleans, November 29, 1974

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TESTING AND TESTS: PEDAGOGICAL VERSUS PUBLIC USES

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Scarvia B. Anderson
John E. Dobbin
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

Scarvia B. Anderson and John E. Dobbin

Educational Testing Service
Atlanta Office
17 Executive Park Dr., N.E.
Atlanta, GA 30329

ED102605

The first typed draft of this speech bore the title TESTING AND TESTS: PEDOLOGICAL VERSUS PUBLIC USES. I put that down to the very messy copy Ms. Prindle had to work from, but, as I read on, I realized that pedological appeared consistently. By the time I got to the last page I couldn't remember what the word was supposed to be. So I consulted a dictionary and realized for the first time (1) that there is a word pedology, (2) that it refers to two rather different sciences, and (3) that I have been engaged in the practice of one of those for many years. Furthermore, I use tests in studies of the development of children. Thus we meet on more common ground than I had suspected we might, for we come from disciplines that share not only a combining form but also a concern for measurement.

The exercise was interesting too in its analogy to the pedagogical uses that a test can have. Just as Ms. Prindle's typescript sent me on a dictionary exploration, so a good test can provoke learning experiences by students. That is a major message I want to leave with you today, although the NCTE program committee has asked me specifically to contrast pedagogical and public uses of tests. My job will be easier if I start by trying to define the two classes of test use:

1. Pedagogical uses of tests and testing cover all those materials and practices in observation of human behavior that are intended to help the teacher do a better job of teaching or the learner do a better job of learning, or both.

CS 201 905

BEST COPY AVAILABLE

2. Public uses of tests and testing include all those materials and practices in observation of human behavior that are intended to help administrators, school boards, legislatures, taxpayers, and other interested groups outside the classroom in evaluation of their educational systems.

One more definition before we get underway. Too many people, teachers and parents alike, act as if they think a "test" is some tool of the devil that leads students to betray their ignorance, reveal the secret contents of their minds, uncover their psyches. It is, of course, nothing of the sort. Every test ever devised for school use is nothing more than a job on which, if he is so inclined, the student can demonstrate his skill, knowledge, or recollection. Asking a student to take a test is exactly the same thing as his mother saying to him, "Show Uncle Albert how you can play the violin." If the kid knows how to play the violin pretty well and is in good health and feels like showing off for Uncle Albert, you can't tell him from Heifetz. But if he hates to play the violin and doesn't play it well, has a stomach ache, and doesn't like Uncle Albert, you can tell him from Heifetz.

Public Uses of Tests and Testing

Remembering that, in spite of the technical jargon that surrounds them, tests are no more than jobs, worksamples, let us consider first the public uses of tests. In the earliest days of public schooling, schools were small, very much a part of their communities, the teachers boarded around with various parents or board members, and what went on in the local school was quite naturally a matter of public knowledge. It was not

unusual to have a visitors' day once a month when students could demonstrate their knowledge and skills for all to see. When school districts and school buildings became larger, some of the intimate knowledge of students' classroom progress disappeared. Finally, in most places, the school district became so large and the state and national governments so deeply involved in education, that parents and other interested citizens found much to discourage them from finding out directly what kind of job the schools were doing. Oh, they were invited to show up at PTA meetings and at the annual open house, but they were enjoined --not always subtly--to "leave education to the professionals."

Now the pendulum is on its return swing. In one legislature after another, representatives of the people are saying to professional educators: "Show us. Prove to us that we are getting quality education of our young people in return for the huge investments of public funds we keep pouring into the schools." The popular term for this movement is "accountability," and, for reasons not entirely clear to me, it has produced a frantic reaction among many educators that is very much like the behavior of a colony of ants suddenly exposed to sunlight.

Fortunately, the wild running about and hiding phase seems to be subsiding now in some places--for example, in Florida and California--and it is possible to see dimly the outlines of what school accountability might be like. Let me describe it briefly, for this is where the public uses of tests and testing comes in.

In the first place, the public (as represented by legislatures) has a somewhat different view of "quality education" from that of most educators. Using information obtained in the ETS survey of education committees in state legislatures, we discover that parents and taxpayers limit their expectations primarily to the "Three Rs." "Give us high school graduates who can read the sorts of material that citizens must read, write legibly and with enough organization to get over a short declarative message, and perform routine computations." When some of the interviewers asked these chairmen of legislative committees about such things as listening comprehension skills, speaking facility, knowledge of the structure of language--they figuratively shrugged their shoulders. "All of that is very well," they said, "provided students learn to read and write."

The implication is clear. If what legislators want students to demonstrate is reading and writing skill, then that is what the tests "for public use" will measure. And, of course, teaching students to read and write has been one of the major businesses of professional teachers of English for decades.

Now, how will such tests of reading and writing be used? In ways that will show individual schools in a bad light? In ways that will endanger the jobs of teachers who have slow or otherwise handicapped classes? Possibly. There may be some legislators or taxpayer groups or even administrators who would like to evaluate single schools and individual teachers on the basis of student performance on statewide or regional tests mandated for accountability purposes, but one hard economic fact may deter them: they can't afford it. Think for a

minute of the material cost, teacher time, and scoring and analysis expense that would be involved in testing every student in every subject in every school every year! This is what those who would like to get a toe-hold on assessment of teachers by the test performance of their students would have to do. Even in states with high investments in education and great interest among legislators, it just about ruins the budget to test a sample of kids in three subjects in three grades out of twelve every year.

So, in spite of the demonic glee that is generated in some quarters over the notion of statewide testing to "measure the quality of our schools and the effectiveness of our teachers," simple economics dictate against such uses. Tests of reading and writing skills can be used in state and systemwide assessment efforts to ascertain the general effectiveness of the whole organization of public education, and I think that few of us would question this as an appropriate use. But they can seldom be used to "judge" single schools and almost never used to evaluate individual teachers.

What should English teachers be expected to do about such public uses of tests? Probably pretty much what you are doing now, but perhaps with a new awareness of the priorities which others place on reading and writing skills in the language arts curriculum. If I were to teach English again--and I did once before I was lured into the less demanding field of psychology or pedology--I would be more inclined to relegate the productive skills to second place in the early grades until children seemed to have a firm command of the receptive skills of reading and writing.

And, at later grades, I would be more willing to abandon the prescribed syllabus in literature with those students who could hardly be expected to read it with understanding, what's more answer essay questions about it, and put my emphasis on those basic skills.

Thus, even though individual teachers cannot be required to take personal responsibility for a statement derived from a public use of tests such as "Thirty percent of the seniors graduating from high schools in City X cannot read or write well enough to function as participating citizens in our society," teachers who work conscientiously and collectively to improve reading and writing skills--not because they're forced to but because they recognize their salience--can take pride in a statement that may appear a few years later to the effect that "Less than twenty percent of the seniors graduating from our high schools are deficient in reading and writing skills."

Some of you may feel that I am putting too much sugar coating on the public uses of testing pill, and I certainly don't deny that tests and test results are sometimes misused and misinterpreted for strictly political purposes. However, an adversary relationship between teachers and politicians certainly will not reduce misuses and abuses. And an understanding relationship might at least stand a chance of doing that.

Pedagogical Uses of Tests and Testing

Now let's turn to the uses of tests and testing to help the teacher teach better and the student learn better. If we can keep thinking of tests as job performances or skill demonstrations, instead of as sneaky peeks into the skulls of students, the process is fairly straightforward.

When students come to us, we inquire into the skills and knowledge they already possess (like the football coach having members of a green squad run through their paces for him). For this reason, we give them tests at the opening of a semester in order that we may teach them some things they don't already know. This is called placement testing--if the job covers many subskills it is sometimes called diagnostic testing--and helps the teacher guess better where to start the instruction of each student. To start a semester or any large unit of instruction without this kind of bench mark measure is sheer folly--like timing a hundred yard dash without knowing when each runner started.

The tests or worksamples that follow the placement testing usually cover small amounts of instruction, so that both the teacher and the student can know without a long wait whether the instruction is "taking" or sliding off without visible effect. Most of these tests will be teacher-made instruments--teacher-made and teacher-scored. Occasionally, we hope, the tests themselves will be learning experiences--or stimulate them, just as a provocative statement in a magazine or a typist's "error" can lead adults to new insights or to seek new information.

Because schools and teachers are the way they are, a great deal of the teaching that goes on represents an effort to get pupils to "know" things. That is, most of the effort is to get the student to remember factual information that ranges from the ridiculously simple to the preposterously complex. There is "content" to be covered in the course. "Remember this because it will appear in the semester exams." No one doubts that a large proportion of pupil learning in school should consist

of remembered information, but teachers too often trap themselves into testing only factual recall even though their goals of instruction are far wider and richer. Being human beings, students almost without exception learn what they know they'll be tested on, and let the rest go by the board. So the teacher who aims instruction toward development of all kinds of insights and appreciations and tastes and critical processes will be bitterly disappointed if he or she routinely limits classroom testing to measures of recall--because the outcomes of that ambitious instruction will be a respectable amount of factual recall among those students who are interested or fearful, no recall among students who don't care, and no provable improvement of anybody in those other goal behaviors: insights and appreciations and tastes and critical processes. So let the test jobs you use to track the progress of students along the path of learning somehow cover all the kinds of learning you hope to have them attain. Let me give you an example, drawn from some work done by the English staff in a large high school, all of whom were committed to teaching Macbeth sometime during the senior year. (See Appendix A.)

For a four-week unit on Macbeth, five teachers who were going to hit the play sometime during the semester got together and wrote down their general goals. These goals ultimately were eight in number: (1) to understand Macbeth in each of several ways, (2) to enjoy Macbeth in each of several ways, (3) to have an increased interest in reading plays and seeing more drama, (4) to develop better taste and preference as

consumers of drama, (5) to know and remember four basic literary aspects of Elizabethan drama found in Macbeth, (6) to have some skill in analyzing the dramatic elements of a play, (7) to recognize, interpret, and retain some human values as expressed in Macbeth, and (8) to appreciate the reasons for Shakespeare's popularity in his time and ours. Now, that is a very sophisticated list of outcomes to seek with a high school class, but these teachers were wise enough not to expect total achievement of all outcomes by all students; they agreed to place proportionately high weights on Goals 1, 5, and 6--all being goals of knowing and remembering factual information--and low but nevertheless present weights on the other goals that had to do with interest and taste and appreciation.

Then these teachers did an interesting thing. For each goal and subgoal in the unit, they wrote down what they called "Symptoms of Learning" (members of my profession have their own term for this, not nearly so clear, which is "behavioral objectives"). These "symptoms" were simply the kinds of things that students might do that could be taken as evidence that they had information or appreciation or interest or taste. To quote a few of the listed symptoms that teachers would accept as evidence that a student was learning to enjoy Macbeth: shows some evidence of emotional reaction to hearing or reading the play, quotes favorite passages voluntarily, seeks and enjoys a part in a dramatization, volunteers opinions, debates points of interpretation, disagrees with the teacher or class, shows special sympathy for a particular character, mimics or burlesques some character--and so on.

Then the last thing these teachers did in preparing their evaluation scheme for the students reading Macbeth was to make an "Assessment of Learning" column in which, for each of the symptoms of learning they had listed, they wrote down one or more ways in which they would observe and record the symptom if it occurred--an observational record for each student. The symptoms of learning that had to do with knowing, understanding, remembering, interpreting, and analyzing, as you might expect, were mostly set up for observation by "tests," multiple choice, short-answer questions, essay questions, and so forth. To observe the symptoms of things such as enjoying and having better taste and appreciating, however, those teachers had to be a little more creative. They relied heavily on performance check lists in which they made an entry each time some student exhibited one or another of the kinds of behavior that the teachers had listed under "symptoms of learning." In order to avoid spending all their time making notes on check lists, these teachers very soon learned to note only the first few instances of a symptomatic behavior--or only those instances that pointed toward some kind of change in the student.

I'd be derelict if I neglected to mention that there are literally hundreds of published tests available to you for pedagogical uses in English. Among them are many tests of great value in observing student growth in reading skills. However, because of the difficulties in specifying scoring procedures and standards, standardized tests of writing ability are scarce. There are, of course, many test tasks that require the examinee to edit or rewrite or find the errors in somebody

else's writing. However, these are not tests of writing ability at all--they are tests of editing, which is something else. So even if it means that you have to read stacks of student papers nearly every night, if you want to find out how each kid writes in order to be able to help him write better you have little choice but to ask him to write for you. You'll do a better job of this if you ask for short papers, let the student decide what he wants to write about, tell him the criteria you are going to use in reviewing and commenting on his paper--then review only on the basis of those previously-announced criteria.

As to measurement of the knowledge and cognitive skills students acquire in the study of literature--the major goals in my illustration of the teaching of Macbeth, for example--I can recommend the Cooperative Literature Tests, which offer matched pairs of tests on each of more than twenty major works of fiction often included in high school English instruction. These tests, build by scholars and teachers of literature, probe the works in considerable depth. It is suggested that students take the first test on their own while they are studying the work. The questions are designed not only to enable students to find out if they are "following" the work fairly well, but also to remind them of points they might have overlooked and of alternate interpretations of characters and events. The second test can be used later as the "final" or part of the "final" on the work.

So I have mentioned some ways in which teachers of English can use standardized tests, teacher-made tests, and systematic observation of pupils in understanding their students' learning needs and watching them change as learning progresses. Informed and sensible use of a variety

of measurement techniques by the teacher does not indicate an over-dependence on testing; rather, it reveals a teacher who knows what she or he wants to do in teaching young human beings and how to tell whether that teaching is working. In the work of a good teacher who is also informed about measurement, teaching and testing and learning can blend into each other so smoothly that is is often not possible to tell where one stops and another starts.

EXHIBIT F

BEST COPY AVAILABLE

GOALS If the unit of instruction on Shakespeare's Macbeth is successful, most students will:	LEARNING EXPERIENCE Student experiences hopefully leading individuals toward goal behavior.	SYMPTOMS OF LEARNING Behavioral evidence that the teacher will accept as a sign that the student is growing toward the stated objectives.	ASSESSMENT OF LEARNING Note: Grades on the Macbeth unit will be assigned as follows: 1) For students in the fast-learning classes a) 75% on achievement of Goals I, V, and VI combined b) 5% each on achievement of Goals II, III, IV, VII, and VIII 2) For students in average and slow-learning classes a) 10% each on achievement of Goals I, V, VI, VII, and VIII b) 50% on achievement of Goals II, III, and IV combined
I. Understand Macbeth - A. As Elizabethan entertainment B. As a classical tragedy C. As a uniquely Shakespearean drama D. As a high-point in the development of the drama from ancient times through church, guilds, Inn Yards, theaters	A. Listening to a recording of Macbeth, reading the textbook, using study guides, hearing lectures, seeing filmed illustrations, reading the play B. Reading the textbook, hearing lectures, using study guides C. Reading the textbook, hearing lectures, seeing filmed illustrations, reading the play D. Reading the textbook, hearing lectures, using the study guide, seeing filmed illustrations	As a result of his learning in the Macbeth unit, the student ..... A. Can describe an Elizabethan theater audience, point out some elements of Macbeth that probably appealed to each segment of that audience, venture an opinion on the "fit" of the play to a modern American TV audience -- naming specifics and recalling some cultural characteristics of Elizabethan England that made Macbeth appeal to those audiences B. Can describe the main dramatic elements that distinguish a classical tragedy -- and illustrate them with references to Macbeth C. Can recount some of the characteristics of Macbeth that identify it as uniquely Shakespearean D. Can recount the principal milestones in development of the drama to the time of Shakespeare	Paper and pencil test - A. Short-answer section. (Example questions) What kinds of people were especially attracted to Shakespeare's plays in his time? Mention one scene from Macbeth that probably was written for each of the kinds of people you named. Which scenes in Macbeth would appeal especially to a TV audience of teenagers? B. Short essay section. (Example) Describe in one paragraph the main dramatic elements of a classical tragedy. Number them. C. Objective or multiple-choice section. (Example) Which of the following dramatic characteristics were unique to Shakespeare in his time? D. Short essay section. (Example) By what means was the drama kept alive and developed until Shakespeare's time?
II. Enjoy Macbeth - A. As a dramatic experience B. As a revelation of Elizabethan thought C. As a link with the heritage of English-speaking people D. As a sample of popular drama	A. Listening to a recording of Macbeth, reading aloud certain passages and dramatizing others B. Reading and discussing passages that are peculiarly Elizabethan in mood or syntax, studying the textbook, hearing lectures C. Locating and discussing passages or ideas pertinent to the history of Anglo-American thought, hearing lectures, reading the text D. Listening to the recording of Macbeth, locating passages with appeal to various kinds of audiences, locating and dramatizing passages with personal appeal, preparing scripts for commercial promotion of the play	A. Shows some evidence of emotional reaction to hearing or reading Macbeth B. Asks questions or contributes comments or seeks additional information on the place of Elizabethan thought in Anglo-American history C. Asks questions or contributes comments or seeks additional information on the character of life in Elizabethan England D. Quotes favorite passages voluntarily, seeks and enjoys a part in a dramatization, does personal research on some aspect of the play or Elizabethan character, volunteers opinions, debates points of interpretation, disagrees with the teacher or class, shows special sympathy for a particular character, mimics or burlesques some character, etc. E. See also evidence listed in III below	Performance checklist - A systematic record noting occurrence and frequency of individual behavior of the kinds noted under "Symptoms", sections II and III
III. Have an increased interest in reading and seeing more drama -		Reads part or all of some other Shakespearean drama, reads part or all of some modern drama, watches all of a Shakespearean drama on TV, mentions opportunities to see or hear serious drama, inquires about such opportunities, comments on dramas he has already seen, brings to class playbills and other artifacts of the drama, suggests dramatizations or other class activities, asks the teacher for more class work in drama, etc., etc.	Performance checklist - A systematic record noting occurrence and frequency of individual behavior of the kinds noted under "Symptoms", sections II and III
IV. Develop better taste and preference as consumers of drama	Comparing Macbeth with various stage, screen, and TV dramas known to most students -- as to plot, character development, dramatic effect, longevity; writing comparative	Comments critically and sensibly on movie or TV drama in comparison with Shakespearean drama, chooses more mature TV drama and avoids or ridicules the less mature, goes to	Performance checklist - A systematic record noting occurrence and frequency of individual behavior of the kinds noted under "Symptoms".

BEST COPY AVAILABLE

V. Know, remember, and recognize in later contacts with drama the following aspects of Elizabethan literary and dramatic art found in Macbeth - A. Dramatic devices 1. Witches, apparitions, the supernatural 2. Trap doors and other stage surprises 3. Bombastic speeches B. Literary devices 1. Blank verse 2. Couplets 3. Images C. Elements of classical drama 1. The classical plot 2. The classical tragic character 3. Dramatic conflict 4. Reversal 5. Recognition D. Drama as a literary form 1. Exposition 2. Rising action 3. Climax 4. Falling action 5. Denouement	A-B. Analyzing the play to locate specific dramatic devices, scanning lines to ascertain verse form, identifying couplets, locating and interpreting images C. Locating in the play the elements of classical tragedy, hearing lectures and seeing illustrations, tracing changes in the relationships of the main characters, graphing the action of the play D. Hearing lectures on the form of the drama, analyzing the structure of Macbeth, graphing the action of the play and labeling the divisions	A. Can recognize Shakespearean dramatic devices in other Shakespearean dramas B. Can recognize Shakespearean literary devices in other Shakespearean dramas, and can recognize and to a limited extent interpret Shakespearean literary images C. Can recall, recount, and identify the main elements of classical drama D. Can ascertain verse form by scanning	Read Act III of Hamlet at home and take the following "open-book" test on it - A. Short-answer section on identification of dramatic devices B. Short-answer section on identification of literary devices C. Essay section requiring interpretation of two given literary images in Hamlet D. Short-answer section on identifying verse forms from Hamlet
VI. Have some skill in analyzing the dramatic elements of a play -	A-B. Analyzing the play to locate specific dramatic devices, scanning lines to ascertain verse form, identifying couplets, locating and interpreting images C. Locating in the play the elements of classical tragedy, hearing lectures and seeing illustrations, tracing changes in the relationships of the main characters, graphing the action of the play D. Hearing lectures on the form of the drama, analyzing the structure of Macbeth, graphing the action of the play and labeling the divisions	A. Can recall the elements of classical tragedy and identifying them in another Shakespearean tragedy B. Can analyze the structure of a Shakespearean tragedy	(open-book test on Hamlet continued) A. Matching-test section identifying the elements of classical tragedy in Hamlet B. One-paragraph essay analysis of Hamlet as a tragic character C. Graph the action of Hamlet
VII. Recognize, interpret, and retain some human values as expressed in Macbeth -	Memorizing self-chosen passages and giving reasons for selection, writing a theme on a self-chosen passage selected as an illustration of an important human value, locating passages with special personal appeal, drawing parallels between Macbeth and some persons in modern history	Can locate in other Shakespearean drama salient passages expressing important human values, can recall value stereotypes from Macbeth, can quote from Macbeth at least two short passages bearing on human values, can interpret from some other Shakespearean drama passages bearing on human values, contributes comparative comments having to do with values	(open-book test on Hamlet continued) A. Short essay section 1) Locate, paraphrase and interpret two passages in Hamlet bearing on human values important to you 2) Quote (loosely) two passages from Macbeth bearing on these same values
VIII. Appreciate the reasons for Shakespeare's popularity in his time and ours -	Discussing the TV production of Macbeth and its rating, discussing the increase in number of Shakespearean festivals, collecting published quotes from Shakespeare in a scrapbook	Comments in class on differences and similarities between Shakespeare's audiences and modern audiences, voluntarily delves into theatrical history, brings theater models and artifacts to class, looks up comments of Shakespeare's contemporaries, reports records of Elizabethan play attendance, etc.	Performance checklist - A systematic record noting occurrence and frequency of individual behavior of the kinds noted under "Symptoms", section VIII